Novel covalent and non-covalent complex-based pharmacophore models of SARS-CoV-2 main protease (M^{pro}) based on molecular docking, molecular dynamics, and MM/PBSA calculations

Yasser Hayek Orduz[1], Andrés Felipe Vásquez[1,3], María Francisca Villegas Torres[2,4], Paola Andrea Caicedo Burbano[4], Luke E. Achenie[5], Andrés Fernando González Barrios[1]

[1]Grupo de Diseño de Productos y Procesos (GDPP), Department of Chemical and Food Engineering, Universidad de los Andes, Bogotá, Colombia.

[2]Centro de Investigaciones Microbiológicas (CIMIC), Department of Biological Sciences, Universidad de los Andes, Bogotá, Colombia.

[3]Naturalius SAS, Bogotá, Colombia.

[4]Grupo Natura, Faculty of Sciences, Universidad ICESI, Cali, Colombia.

[5]Department of Chemical Engineering, Polytechnic Institute and State University from Virginia (Virginia Tech), USA.

Presenting author: Yasser Hayek Orduz, y.hayek10@uniandes.edu.co

One of the most promising proteins to inhibit the replication cycle of SARS-CoV-2 and, hence, decrease its viral load, is the nonstructural protein 5 (NSP5), also known as 3-chymotrypsin-like protease ($3CL^{pro}$) or main protease (M^{pro}). Main protease inhibitors can be categorized into two types: covalent inhibitors and non-covalent inhibitors. However, and although there are studies on the dynamic behavior or the steric and electronic features of covalent and non-covalent inhibitors in complex with main protease, the differences in behavior between covalent and non-covalent inhibitors in virtual screening processes are not yet known. Furthermore, protein-ligand interaction studies have been done using pharmacophore modeling but no models have been designed taking into account the active site flexibility.

In this research work, with the aim of modeling, identifying, and discriminating the essential physicochemical features that favor the interaction between SARS-CoV-2 M^{pro} and either covalent and non-covalent inhibitors, we designed novel complex-based pharmacophore models, using a cascade protocol of docking, molecular dynamics (MD), k-means clustering, principal component analysis and Molecular Mechanics Poisson-Boltzmann Surface Area (MM/PBSA) free energy calculations. At first, we selected 15 inhibitors with experimental pose and 38 inhibitors without experimental pose but with experimental inhibitory activity. The 38 inhibitors underwent a flexible molecular docking

process to calculate predicted poses in the active site. Subsequently, the poses of all inhibitors were used to carry out 1000 ns MD simulations. MD trajectories were employed for a MM/PBSA calculation of protein-ligand binding free energy and an affinity ranking was generated based on free energy data. Finally, the MD trajectory of the complexes located in the first positions of the affinity ranking was used to design pharmacophoric models, that consider active site flexibility, using principal component analysis (PCA) and k-means clustering.

Among the 53 protein-ligand complexes we found 23 inhibitors that showed high affinity towards SARS-CoV-2 M^{pro}. We noticed that the chemical structure of most identified active ligands was based on peptidomimetics. The MD trajectories of the 23 protein-ligand complexes were used in conjunction with clustering techniques to obtain 18 active site conformations. By a visual inspection of the conformations, appreciable changes were detected on the surfaces of subsites S2 and S4. It may suggest that active site flexibility plays an important role in protein-ligand interactions. Additionally, the clustering data was employed to generate 36 complex-based pharmacophoric models divided into 18 covalent and 18 non-covalent models. The pharmacophore model validation showed two models with 0.93 and 0.73 ROC-AUC values, respectively, for the covalent and non-covalent categories. Since the models were created based on conformational data, they take into consideration active site flexibility. Furthermore, the high ROC-AUC values indicate that these models have a high ability to predict potential protease inhibitory compounds. A further benefit of using pharmacophore models is the possibility of discovering new scaffolds that may act similarly to peptidomimetics and exhibit high affinities towards SARS-CoV-2 M^{pro} (scaffold hopping). We expect to use our pharmacophore models in the design of new molecules that may function as a treatment for COVID-19.